Human Immunology 77 (2016) 313–316

Contents lists available at ScienceDirect

journal homepage: www.elsevier.com/locate/humimm

Rapid Communication

Identification and characterization of novel HLA alleles: Utility of next-generation sequencing methods

Nicholas K. Brown, Taba Kheradmand [1], Jinguo Wang, Susana R. Marino *

Department of Pathology, University of Chicago Medicine, Chicago, IL, United States

ARTICLE INFO

Article history:
Received 10 July 2015
Revised 22 December 2015
Accepted 3 January 2016
Available online 4 January 2016

Keywords:
Novel HLA alleles
SBT
NGS
Sanger sequencing
Next-generation sequencing
Allele-specific primers

ABSTRACT

The HLA genes are the most polymorphic of the human genome, and novel HLA alleles are continuously identified, often by clinical Sanger sequencing-based typing (SBT) assays. Introduction of next-generation sequencing (NGS) technologies for clinical HLA typing may significantly improve this process. Here we compare four cases of novel HLA alleles identified and characterized by both SBT and NGS. The tested NGS system sequenced broader regions of the HLA loci, and identified novel polymorphisms undetected by SBT. Subsequent characterization of the novel alleles in isolation of coencoded alleles by SBT required custom-designed primers, while the NGS system was able to sequence both alleles in phase. However, the tested assay was unable to amplify buccal cell DNA for subsequent NGS sequencing, presumably due to the lower quality of these samples. While NGS assays will undoubtedly increase novel allele identification, more stringent DNA sample requirements may be necessary for this new technology.

1. Introduction

The HLA gene complex is the most polymorphic region of the human genome [1], which continues to challenge clinical laboratories tasked with typing these genes. The widespread adoption of molecular techniques, including Sanger sequencing-based typing (SBT), has allowed the routine reporting of HLA genes at high resolution, which includes a set of alleles encoding for identical protein sequences at the peptide-binding regions [2]. Due to high polymorphism of the HLA genes and the inability of the SBT technique to determine phasing of sequences, complimentary methods are often required to obtain high resolution typing. Lack of phase information also makes clinical SBT assays ill-suited to characterize novel alleles, as such alleles need to be sequenced in isolation of the other encoded allele [3]. The introduction of next-generation sequencing (NGS) technologies, which provide clonal sequence information and may be used to determine phase over long stretches of DNA, has the potential to overcome many of the limitations of SBT.

Herein we describe four cases of novel sequences identified in DNA obtained from peripheral blood lymphocytes (PBL) of hematopoietic cell transplant (HCT) patients undergoing treatment at the University of Chicago Medicine. These cases were identified by SBT during routine clinical testing, and were further characterized by SBT and/or NGS technologies. The overall process of novel allele identification and sequencing by both SBT and NGS are compared.

2. Materials and methods

2.1. Patient demographics

Between May 2013 and January 2014, 354 patients and donors were typed at high resolution in support of the HCT program, and four patient DNA specimens from PBL were identified with putative novel HLA alleles. High-resolution typing was performed using AlleleSEQR SBT kits from Celera. Exons 2–4 (HLA-A, B and C), 2 and 3 (DQB1) or 2 (DRB1) were sequenced, and data were analyzed by SBTEngine software (GenDX). Genomic DNA from peripheral blood was used for initial typing. As tumor cells may harbor somatic mutations [4–7], it is our laboratory's policy to use a buccal cell specimen to confirm germline encoding of any putative novel allele.

* Corresponding author at: Transplant Immunology and Immunogenetics Laboratory, Department of Pathology, University of Chicago Medicine, MC 0006, 5841 S. Maryland Ave., Chicago, IL 60637-1470, United States.
E-mail address: smarino@bsd.uchicago.edu (S.R. Marino).
[1] Present address: Gift of Life Michigan, 3861 Research Park Dr., Ann Arbor, MI 48108, United States.

http://dx.doi.org/10.1016/j.humimm.2016.01.001
0198-8859/

Table 1
Primers used for PCR amplification.

Case	Locus amplified	Encoded alleles	Exon	Primer sequence[a]
1	A	*66 novel, *69:01	2	F: GAACCTGGCCCGGACCC[T]GA R: CCCAGACCTGGGCGGGTGAG
			3	F: TTAGGCCAAAAATCCCC[T]CG R: GGTGTGTGTCAAAGGGAACC
2	B	*07 novel, *52:01	2	F: AGACCTGGGCCGGTGAGTG R: ACGTCGCAGCCGTACATGC
			3	F: ACGCCGCGAGTCCGAGAGA R: CCATTCAACGGAGGGCGACA
			4	F: TGGAGAACGGGAAGGACAA R: AAGACGGCTCTGGGAAAGG
	DQB1	*06:01 novel, *05:01	1	F: GCAATTTTCTCTCCCCTGAA R: AGACATCATTGCTGCCTTGT
			2	F: TCAGGCAGGGACAGGGCTTT R: AGGCAAACGCACAAGGCAGA
			3	F: GCGGGGAAAGATGTGTCATT R: TGCCATGGAGCAAGAGATGT
			4	F: TGGGAATCACAGAAGACTAGACA R: CACTCTCTCATCCCAAAGGAA

[a] Underlined nucleotides indicate polymorphisms compared to the other encoded allele; boxed nucleotides indicate introduced mismatches compared to both encoded alleles.

Table 2
Primers used for Sanger sequencing.

Case	Sequenced amplicon	Exon	Primer sequence
1	A*66 novel	2	F: AAAATGAAACCGGGTAAAGG R: AACCCTCGTCCTGCTACTCTC
		3	F: AAAAATCCCCTCGGGTTG R: GTCTTCCCGTCATTCCTTCA
2	B*07 novel	2	F: AlleleSEQR primer R: AlleleSEQR primer
		3	F: GAGATCCGCCTCCCTGAG, AlleleSEQR primer R: TTCTAGCGCTGATCCCATTT, AlleleSEQR primer
		4	F: AlleleSEQR primer R: AlleleSEQR primer
	DQB1*06:01 novel	1	F: TCCCTGAGGCACTCATTCTT R: CCAATGCAGGATCCCATAAT
		2	F: TGGCTGTAGGAAGGCAGATT R: GAAACCTGCAGAGCAGAGGA
		3	F: GCAGGGACCATGGTATGTTG R: CTTCTGTGATTCCCAGCTCAG
		4	F: CTTACGGGTGGTCTGAATGG R: TGCCTTTTCCAATCTTGTCC

2.2. Sanger sequencing of novel alleles

Allele-specific (AS) primers were designed to amplify genomic regions about 800–1000 base pairs in length, containing an exon of the putative novel allele, separately from the other encoded allele. Primers taking advantage of single nucleotide polymorphisms were designed according to a previously-published method [8]. Briefly, primers were designed so that the 3′ terminal nucleotide aligned to the polymorphism. An additional nucleotide difference was inserted at the third position from the 3′ terminus, with the identity of the introduced mismatch based on empirical data [8]. For sequences that had more than one polymorphism closely spaced, primers were designed to take advantage of these by placing the mismatches at the 3′ end. Combinations of Exons 1, 2, 3, and 4 were PCR amplified using Taq polymerase (Celera) and standard PCR buffer (Applied Biosystems). Table 1 lists all amplification primers used. Primers for sequencing were designed to be internal to the amplification primers, and the amplicons were sequenced from each direction using BigDye Terminator v1.1 Cycle Sequencing Kit (Life Technologies). For one novel allele, B*07, sequencing primers from the AlleleSEQR clinical SBT kit (described above) were used. Table 2 lists all amplification primers used. Paired sequence reads were analyzed, and consensus sequences were determined, using SeqTrace software [9].

2.3. Next-generation sequencing

The Holotype X4 sequencing kit (Omixon) was used to amplify the *HLA-A*, *B*, *C*, *DRB1* and *DQB1* loci. Amplicons were fragmented enzymatically and "barcode" adapters were ligated on both a